ALGEBRAIC CURVES

Christoph M. Hoffmann

CSD-TR-675
May 1987

Algebraic Curves

C. M. Hoffmann*
Computer Science Department
Purdue University

May 12, 1987

Abstract

We consider the problem of tracing algebraic curves by computer, using a numerical technique augmented by symbolic computations. In particular, all singularities are analyzed correctly. The methods presented find application in solid modeling and robotics.

1 Introduction

In this paper we discuss preliminary results for tracing algebraic curves. Planar algebraic curves of the form $f(x, y) = 0$ are considered, as are space curves that are the intersection of two algebraic surfaces, $f(x, y, z) = 0$ and $g(x, y, z) = 0$. The scenario is as follows: We are given a point p on the curve, and a direction of traversal. We wish to trace succeeding curve points on the same branch, and we would like to trace them through singularities.

A reliable solution to this problem has immediate applications to solid modeling and geometric design. For example, the well-known Boolean operations on solids require tracing surface intersections, for the purpose of determining the surface of the resulting solid [9]. Here, algebraic space curves arise when the intersecting solids are bounded by algebraic faces. If the

*Supported in part by the Office of Naval Research under contract N00014-86-K-0465 and the Institute for Mathematics and its Applications at the University of Minnesota.

surface is composed of parametric patches, e.g., rational B-spline surfaces, then planar algebraic curves can be obtained [6].

Since an extensive amount of curve tracing is required for each modeling operation, it is advisable to pay attention to efficiency as well as reliability. For this reason, precise methods such as the cylindrical algebraic decomposition due to Collins and its variants [4,11] have not been considered here. This does not imply that these presently very compute intensive methods must remain of theoretical interest only, but significantly more work is required before we can accurately gauge whether in the specialized context of solid modeling versions of these algorithms exist that can be used without serious efficiency degradations.

As point of departure, we use a straightforward numerical method that approximates the curve locally by its truncated Taylor series, and then performs a Newton iteration to correct the accumulated error. This approximation has many interesting properties. For instance, the Taylor series is a special case of the Puiseux series that can be used to approximate the curve at singular points as well as at regular points. Thus, there is a basis for developing a uniform framework for studying approximations to algebraic curves. Note, however, that there are alternative curve approximations based on special classes of polynomials that offer different advantages, [7], and more study is needed before the relative merits can be fully appreciated.

In many cases the numerical procedure suffices and copes acceptably well with certain singularities, e.g., with normal crossings and with tacnodes that are not very complicated. It is not fully reliable, however, and will fail at cuspidal singularities. For this reason it is augmented by a mapping technique that exploits the fact that by a suitable birational map any singularity can be resolved. That is, such a map will transform the singular point into one or more nonsingular ones, while not creating new singularities. Fortunately, suitable maps can be found easily in the planar case. In the space curve case the situation is not so simple, and more work is required to find attractive algorithms.

Even in the planar case a number of details must be addressed before curve desingularization can be automated. These include finding reliably the locus of the singularity, controlling numerical inaccuracies that arise from the various desingularization maps, and establishing the correct correspondence of orientation between the curve and its transformations.

We structure this paper as follows: After explaining concepts and nota-

tion in Section 2, we devote Section 3 to a description of the simple numerical procedure for following curves. In Section 4, we explain the correspondence between the Taylor series and the notion of a place of a curve, which yields a straightforward method for extending the numerical procedure so as to cope with low order singularities consistently. The extension is of limited value, however, as it involves solving systems of polynomial equations whose degree depends on the order of the singularity analyzed. Section 5 concentrates on desingularizing planar algebraic curves and discusses how problems such as orientation correspondence can be solved. Examples of various types of singularities are also given. Section 6 then discusses the desingularization of space curves. Sections 3, 4 and 5 summarize work reported in [2,8].

2 Concepts and Notation

We consider algebraic space curves given as the intersection of two algebraic surfaces $f(x,y,z)=0$ and $g(x,y,z)=0$, where f and g are polynomials in x, y, and z. This is not the most general definition of space curves; certain curves require intersecting more than two surfaces in order to exclude extraneous components. The partial derivatives of f are written by subscripting, e.g., $f_x = \partial f/\partial x$, $f_{xy} = \partial^2 f/(\partial x \partial y)$, and so on. Since f and g are analytic, $f_{xy} = f_{yx}$ etc.

Vectors and vector functions are denoted by bold letters. The *inner product* of two vectors $\mathbf{a}$ and $\mathbf{b}$ is the scalar $\mathbf{a}\cdot\mathbf{b}$. The *length* of the vector $\mathbf{a}$ is $|\mathbf{a}| = \sqrt{\mathbf{a}\cdot\mathbf{a}}$. The *cross product* of the vectors is the vector $\mathbf{a}\times\mathbf{b}$.

The *gradient* of the surface f at the point $p=(x,y,z)$ is the vector $\nabla f = (f_x, f_y, f_z)$, where the partials are evaluated at p. If not zero, it is a vector that is normal to the surface at p. A point $p=(x,y,z)$ is *regular* on f if the gradient of f at p is not null; otherwise the point is *singular*. The *Hessian* of f at point p is the tensor

$$H_f = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix}$$

where the partials are evaluated at p.

The intersection curve of f and g is denoted $\mathbf{r}(s)$ and is considered a vector function of the argument s. In Section 3, we will determine an approximation of $\mathbf{r}$ in which s is the arc length measured from some initial point on the curve. As usual, the derivatives of $\mathbf{r}(s)$ are denoted $\mathbf{r}'$, $\mathbf{r}''\ldots\mathbf{r}^{(m)}$.

A point p of the intersection curve $\mathbf{r}(s)$ is *regular* if p is regular on both f and g and if the gradients ∇f and ∇g are linearly independent. That is, the surfaces are not singular at p and intersect transversally. A point p is *singular* on $\mathbf{r}(s)$ for one of the following reasons:

1. The gradients ∇f and ∇g are nonzero and linearly dependent.
2. One of the gradients, say ∇g is zero, but the other is not.
3. The gradients ∇f and ∇g are both zero.

We note that Cases 1 and 2 do not differ in substance.

The *initial form* of f is the polynomial formed by all terms of lowest order in f. For example, the initial form of $x^2 - 2x + y^2 + z^2$ is $-2x$. The initial form approximates the surface in the neighborhood of the origin. In the above example, the surface "looks like" the plane $x = 0$ near the origin. This plane is the tangent plane to f at the origin. In particular, if p is the origin, then p is regular on f if the initial form of f is linear. Otherwise p is singular.

At each point $p = \mathbf{r}(s)$ of a space curve, an intrinsic coordinate system is provided by the *orthonormal triad*. The triad consists of three perpendicular unit vectors, namely the tangent $\mathbf{t}$, the principal normal $\mathbf{h}$, and the binormal $\mathbf{b}$. The Frenet-Serret formulas relate the triad to arc length, curvature, and torsion of the curve at p. With s the arc length, ρ the radius of curvature, and τ the radius of torsion we have

$$\frac{d\mathbf{t}}{ds} = \frac{1}{\rho}\mathbf{h}, \qquad \frac{d\mathbf{b}}{ds} = -\frac{1}{\tau}\mathbf{h}, \qquad \frac{d\mathbf{h}}{ds} = \frac{1}{\tau}\mathbf{b} - \frac{1}{\rho}\mathbf{t}.$$

A planar algebraic curve is given by its implicit equation $f(x, y) = 0$, f a polynomial in x and y. Equivalently, we can think of the curve as the intersection of the surfaces f and $z = 0$. Note that f is then a cylinder with the base line $f(x, y) = 0$. All concepts explained above can therefore be transferred to the planar case.

3 Numerical Tracing

For space curves, the simplest situation arises when tracing the curve in a neighborhood in which both surfaces are nonsingular and intersect each

other transversally. This means that the gradients of f and g do not vanish along the curve and are linearly independent vectors. In this case we formulate a system of linear equations from which to obtain the local approximation to the intersection curve at the point p.

3.1 The Basic Method for Space Curves

Since the curve $\mathbf{r}(s)$ must satisfy both f and g identically, each coefficient in the Taylor series of $f(\mathbf{r}(s))$ and $g(\mathbf{r}(s))$ will be zero. Let $p = \mathbf{r}(0)$ be a point on the intersection. Then

$$f(\mathbf{r}(s)) = f(0) + s\nabla f \cdot \mathbf{r}'(0) + \frac{s^2}{2}[\nabla f \cdot \mathbf{r}''(0) + \mathbf{r}'(0) \cdot H_f \cdot \mathbf{r}'(0)] + \cdots$$

and similarly,

$$g(\mathbf{r}(s)) = g(0) + s\nabla g \cdot \mathbf{r}'(0) + \frac{s^2}{2}[\nabla g \cdot \mathbf{r}''(0) + \mathbf{r}'(0) \cdot H_g \cdot \mathbf{r}'(0)] + \cdots$$

This leads to the following system of equations, for $m = 1, 2, \ldots$:

$$\begin{aligned} \nabla f(p) \cdot \mathbf{r}^{(m)}(0) &= B_{1,m} \\ \nabla g(p) \cdot \mathbf{r}^{(m)}(0) &= B_{2,m} \end{aligned} \tag{1}$$

The quantities $B_{1,m}$ and $B_{2,m}$ are expressions in the partial derivatives of f and g and the lower-order derivatives of $\mathbf{r}$. For $m = 1$ we have

$$B_{1,1} = B_{2,1} = 0$$

and for $m = 2$

$$\begin{aligned} B_{1,2} &= -\mathbf{r}' \cdot H_f \cdot \mathbf{r}' \\ B_{2,2} &= -\mathbf{r}' \cdot H_g \cdot \mathbf{r}' \end{aligned}$$

For higher values of m the expressions $B_{i,m}$ are more complex.

With the assumption of independent gradients at $\mathbf{r}(0)$, the solution to the system has the form

$$\alpha_m \nabla f + \beta_m \nabla g + \gamma_m \nabla f \times \nabla g$$

γ_m can be chosen arbitrarily, and the other coefficients satisfy the nonsingular system

$$\begin{pmatrix} \nabla f \cdot \nabla f & \nabla f \cdot \nabla g \\ \nabla g \cdot \nabla f & \nabla g \cdot \nabla g \end{pmatrix} \begin{pmatrix} \alpha_m \\ \beta_m \end{pmatrix} = \begin{pmatrix} B_{1,m} \\ B_{2,m} \end{pmatrix} \tag{2}$$

System (1) is underdetermined. When using it to approximate the intersection curve, the choices of γ_m relate to the ability to obtain equivalent approximations, e.g., parameterized by cs instead of s, where $c \neq 0$ is a constant. By using the Frenet-Serret formulas, we can interpret these choices. We set $\mathbf{r}' = d\mathbf{r}/ds = \mathbf{t}$, to obtain

$$\mathbf{r}'' = \frac{1}{\rho}\mathbf{h}, \qquad \mathbf{r}''' = \frac{1}{\rho\tau}\mathbf{b} - \frac{1}{\rho^2}\mathbf{t} - \frac{d\rho/ds}{\rho^2}\mathbf{h}$$

For $m = 1$ we have $B_{1,1} = B_{2,1} = 0$, hence $\alpha_1 = \beta_1 = 0$. We let

$$\mathbf{r}' = \frac{\nabla f \times \nabla g}{|\nabla f \times \nabla g|}$$

be the solution of the system, so that the parameter s corresponds to the arc length of the intersection curve.

For $m = 2$ we choose $\gamma_2 = 0$. This implies that $\mathbf{r}''$ is orthogonal to $\mathbf{r}'$, and that from the system solution $\mathbf{r}'' = \alpha_2 \nabla f + \beta_2 \nabla g = \mathbf{h}/\rho$ both the principal normal $\mathbf{h}$ and the radius of curvature ρ is determined.

For $m = 3$ we get $\gamma_3 = -(\mathbf{r}'' \cdot \mathbf{r}'')$, since both $\mathbf{b}$ and $\mathbf{h}$ are orthogonal to $\mathbf{t}$. This determines both $\mathbf{r}'''$ and τ.

The curve approximation so determined can be used in a neighborhood of $\mathbf{r}(0)$ whose size can be estimated by the magnitude of the second and third order terms. If the ratio of $\delta^3|\mathbf{r}'''|/6$ to $|\mathbf{r} + \delta\mathbf{r}' + \delta^2\mathbf{r}''/2|$ is small, then the higher order derivatives contribute little and we have not deviated too much from the true intersection. By halving or doubling a standard distance repeatedly, the stepping size can be adjusted according to curvature and torsion. It is necessary to establish a minimum stepping size since near-singular curves can have areas of arbitrarily high curvature where repeated halving might lead to unacceptable running times.

At the end of the current approximation to $\mathbf{r}$ a point P_0 is reached that is near the curve of intersection but not on it. Beginning with this point, we find a sequence of points $P_1, P_2, \ldots$ that converges to a point p on the curve, using Newton's method. With $P_{k+1} = P_k + \Delta_k$, we want to solve the system

$$\begin{aligned} \nabla f(P_k) \cdot \Delta_k &= -f(P_k) \\ \nabla g(P_k) \cdot \Delta_k &= -g(P_k) \end{aligned} \tag{3}$$

to obtain $\Delta_k \approx sP'_k$. Assuming linearly independent gradients,

$$P_k = \alpha_k \nabla f + \beta_k \nabla g + \gamma_k \mathbf{t},$$

where $\mathbf{t}$ is orthogonal to both gradients evaluated at P_k. Since a change in the direction of $\mathbf{t}$ does not change the values of f and g appreciably, we set $\gamma_k = 0$, thereby obtaining a unique solution for Δ_k. We then set $P_{k+1} = P_k + \Delta_k$.

After the point p is found with acceptable accuracy, a new approximation of $\mathbf{r}(s)$ centered at p is determined.

3.2 The Planar Case

The planar curve $f(x, y) = 0$ arises as intersection of the cylinder $f(x, y) = 0$ and the plane $z = 0$. It can be traced this way as a space curve. In the system (1) the second equation specializes to $\mathbf{r}_z^{(m)} = 0$. Here $\mathbf{r}_z^{(m)}$ denotes the z component of the m^{th} derivative of $\mathbf{r}$. Moreover, since all partial derivatives by z of f are zero, the first equation takes the form

$$f_x \mathbf{r}_x^{(m)} + f_y \mathbf{r}_y^{(m)} = C_m$$

Thus, there is no difference between considering the intersection curve $\mathbf{r}$ or the planar curve.

As before, choosing $\gamma = 0$ for the Newton iteration means that we approach the curve along the local normal direction. An implementation could be specialized, but there appears to be no significant penalty for tracing the curve in space.

3.3 Implementation

The numerical tracing procedure has been implemented in Fortran by R. Lynch on a VAX 8600. With minor modifications, it has then been ported to a Symbolics Lisp machine. Figures 3.1 through 3.4 show some examples of curve traces so obtained. The planar curves have been traced as the intersection of $f(x, y) = 0$ with $z = 0$.

In our experience, nodal singularities cause no problems as long as the tangent directions of the intersecting branches are sufficiently separated. Since the curve orientation may reverse at singularities (c.f. Subsection 5.4 below), the tracing program must be augmented so as to maintain consistent

tangent direction. However, the program cannot trace through cuspidal singularities. Many tacnodes are handled reliably, but inflections at the singularity are not recognized. Thus both the curve $f_1 = y^2 - x^4 - y^4 = 0$, shown in Figure 3.3, and the curve $f_2 = y^2 - x^6 - y^6 = 0$, shown in Figure 3.4, are traced as if they have two real components tangentially meeting at the origin. While this is correct for f_1, it is not correct for f_2 which consists of a single real component with the two branches at the origin each having a point of inflection.

4 Algebraic Extensions of the Method

We derived the equations (1) based on the assumption that the two surfaces intersect transversally and are not singular at the point $\mathbf{r}(0)$ of interest. Clearly, the radius of convergence of the power series about such a point cannot include any singular points of the curve. Nonetheless, the system of equations remains valid even when we are at a singular curve point. The reason for this is that Taylor's theorem is a special case of more general theorems.

Informally, a *place* of the planar curve $f(x, y) = 0$ is a pair of power series

$$x(s) = \sum_{k \geq 0} a_k s^k, \qquad y(s) = \sum_{k \geq 0} b_k s^k$$

such that $f(x(s), y(s))$ is identically zero. The *center* of a place is the point $(x(0), y(0))$ on the curve. Newton's Theorem states that centered at every curve point there is at least one place of f. Likewise, we define a place of the space curve $\mathbf{r}$ as the triple $(x(s), y(s), z(s))$. Since every space curve is the birational image of a planar curve [13,14], there is at least one place centered at every point of the space curve.

The connection between the notion of place and the equation system (1) in the previous section is established as follows. Centered at the point p we consider the place

$$\mathbf{r}(s) = \sum_{k \geq 0} (a_k, b_k, c_k) s^k$$

where $p = \mathbf{r}(0)$. The derivative of the place is defined by

$$\mathbf{r}'(s) = \sum_{k \geq 1} (a_k, b_k, c_k) k s^{k-1}$$

Higher order derivatives are defined analogously.

Since p is assumed to be on the intersection curve of f and g, we know that $f(\mathbf{r}(s))$ and $g(\mathbf{r}(s))$ are identically zero, from which a system of equations is obtained for $m = 1, 2, 3, \ldots$

$$\begin{aligned} K_{1,m} &= 0 \\ K_{2,m} &= 0 \end{aligned} \tag{4}$$

Here $K_{1,m}$ is the coefficient of s^m in the power series $f(\mathbf{r}(s))$ and $K_{2,m}$ is the coefficient of s^m in the power series $g(\mathbf{r}(s))$. This leads to the following

Theorem *For all $m \geq 1$ the equation $\nabla f \mathbf{r}^{(m)}(0) = B_{1,m}$ is equivalent to the equation $m! K_{1,m} = 0$.*

The analogous statement holds for g and $K_{2,m}$. The proof is by induction on the terms of f; see [8] for details. How the series are obtained from system (4) is illustrated by an example.

Consider the cylinders $f = x^2 + y^2 + 2x = 0$ and $g = x^2 + z^2 + 4x = 0$. Their intersection is an irreducible space curve of degree 4 with a singular point at the origin. At the singular point we have the following equations:

$$\begin{aligned} a_1 &= 0 \\ a_1 &= 0 \\ a_1^2 + 2a_2 + b_1^2 &= 0 \\ c_1^2 + 4a_2 + a_1^2 &= 0 \\ 2a_1a_2 + 2a_3 + 2b_1b_2 &= 0 \\ 2a_1a_2 + 4a_3 + 2c_1c_2 &= 0 \\ a_2^2 + 2a_1a_3 + 2a_4 + b_2^2 + 2b_1b_3 &= 0 \\ a_2^2 + 2a_1a_3 + 4a_4 + c_2^2 + 2c_1c_3 &= 0 \\ &\vdots \end{aligned}$$

One of the solutions to this system is

$$\begin{aligned} x(s) &= -s^2 \\ y(s) &= \sqrt{2}s - \frac{1}{2\sqrt{2}}s^3 \cdots \\ z(s) &= 2s - \frac{1}{4}s^3 \cdots \end{aligned}$$

In principle, this approach can be used to extend the tracing method of Section 3 so as to handle singular points, but it may become computationally expensive. Efficient strategies for solving these equations may exist. For example, one can always choose the coefficients of one series, say for $x(s)$, such that $|a_k| = 1$ for one specific k, and all other coefficients are zero [14].

5 Planar Algebraic Curves

The problem of tracing a curve reliably through any singularity is partially solved by the following theorem from algebraic geometry, e.g. [1]:

Theorem *Any given algebraic plane curve can be transformed, by a birational transformation, into a curve devoid of singularities.*

The proof proceeds by an inductive argument that builds up the required birational transformation through a sequence of elementary, quadratic transformations. It is easy to understand that these transformations resolve ordinary singularities, but how progress is made on irregular singularities is more subtle, and we will not discuss it here. Different proofs of the theorem are found in, e.g., [1,10,13,14]. We restrict our attention to those transformation properties that are needed in order to understand how to derive an algorithm from the theorem.

5.1 Desingularization

We map a planar curve $f(x, y) = 0$ to a curve $g(x_1, y_1) = 0$ by the quadratic transformation

$$x_1 = x \qquad y_1 = \frac{y}{x} \tag{5}$$

If f has a singular point of order m at the origin, then

$$f(x_1, x_1y_1) = x_1^m g(x_1, y_1)$$

We call $f(x_1, x_1y_1)$ the *total transform*, and $g(x_1, y_1)$ the *proper transform* of f. The y_1-axis is called the *exceptional line*. Figures 5.1 and 5.2 show two examples of a curve f and its proper transform.

Intuitively speaking, applying the quadratic transformation separates intersecting curve branches that have different tangent directions. To appreciate this, note that the line $y - mx = 0$ is transformed to the line

$y_1 - m = 0$. Moreover, all points (x, y) of the x-y plane with $x \neq 0$ are in $1-1$ correspondence with points $(x, y/x)$ of the x_1-y_1 plane. A point $(0, y)$ of the x-y plane with $y \neq 0$ is mapped to infinity in the x_1-y_1 plane, and the origin of the x-y is mapped to the exceptional line. The effect is that the singular point is "blown up" to the line $x_1 = 0$, and that the branches of f at the origin are separated or, in a precise sense, made less singular. The proof of the theorem shows that after a finite number of quadratic transformations all singularities are removed. We wish to trace through a singular point as follows:

1. We trace the curve f using the basic method of Section 3.

2. When approaching a singularity, we notice at some point p that the determinant of the system becomes too small. We then locate the singular point as described below, and move it to the origin by translating the coordinate system.

3. Now the quadratic transformation is applied yielding the proper transform g.

4. We traverse g beginning at the point p_1 corresponding to p, until we are past the singularity of f and the system determinant is large enough to continue traversing f accurately.

Note that we may have to traverse recursively iterated transforms of f, since the applied quadratic transform may not have fully desingularized the corresponding branch of g. Moreover, care must be exercised in correlating the orientation of g and of f to maintain proper traversal direction.

5.2 Locating the Singularity

When approaching a singular point p, the partial derivatives f_x and f_y of f vanish. If the singularity has higher order, then higher order partial derivatives also vanish.

The singular point is defined as the intersection of the curves $f = 0$, $f_x = 0$, and $f_y = 0$. When traversing the curve f, we have approached the singularity to a point P_0 for which the partial f_x and f_y drop in value below a threshold μ. We use a Newton iteration to construct a sequence of point approximations P_i that converges to the singularity. The iteration is

governed by the following system:

$$\begin{pmatrix} f_x(P_i) & f_y(P_i) \\ f_{xx}(P_i) & f_{xy}(P_i) \\ f_{xy}(P_i) & f_{yy}(P_i) \end{pmatrix} \begin{pmatrix} \delta_x \\ \delta_y \end{pmatrix} = - \begin{pmatrix} f(P_i) \\ f_x(P_i) \\ f_y(P_i) \end{pmatrix}$$

whose solution determines the next approximation to the intersection as $P_{i+1} = P_i + (\delta_x, \delta_y)$. Since this system is overconstrained, we solve it using the least-squares method by solving the 2×2 system

$$A^{\mathrm{T}} A \Delta = -A^{\mathrm{T}} b \tag{6}$$

where A is the 2×3 matrix, $\Delta = (\delta_x, \delta_y)$, and b the righthand side vector.

If the singularity has higher order, the system (6) is singular. In this case we determine which higher order partials also vanish. For each vanishing higher order partial derivative h of f, the matrix A is augmented by the row $(h_x(P_i), h_y(P_i))$ and the vector b by the entry $h(P_i)$. This process continues until $D = A^{\mathrm{T}} A$ has full rank. For exceptional values it is possible that D has rank 1 or zero even though no additional partials of f vanish. This means that we happen to approach the singular point crossing a specific algebraic curve given by the symbolic determinant of D. In this case, a random perturbation of the point should correct the problem. So far, we have not encountered this problem in practice.

5.3 Passing to the Transformed Curve

By a translation of f, the coordinate system is centered at the singular point q just found. This may introduce spurious terms that are controlled based on the information obtained during the iteration locating q. Recall that the vector b in the iteration contains all partial derivatives that vanish. Consequently, if the partial h appears in b, then the corresponding monomial term must be absent in the translated curve. For example, let $\tilde{f}$ be the translation of f, and assume that b contains the vanishing partials f_x, f_y, f_{xx}, f_{yy}, f_{xy}, f_{xyy}, and f_{yyy}. Then $\tilde{f}$ must not contain the terms x, y, x^2, y^2, xy, xy^2, and y^3. Shoulo such terms appear in $\tilde{f}$ with small coefficients, due to numerical imprecision, they are now removed.

Having centered the singularity at the origin, we apply the quadratic transformation (5). Since this transformation maps the line $x = 0$ to infinity, the branch of $\tilde{f}$ we traverse must not have the y-axis as tangent. If it does,

we rotate $\tilde{f}$ by

$$x' = y \qquad y' = -x$$

before applying the quadratic transformation.

In practice, one applies instead the quadratic transformation

$$x_1 = \frac{x}{y} \qquad y_1 = y$$

in which the x-axis becomes the exceptional line. Which quadratic transformation is used is decided based on the current tangent direction of f in the traversal.

5.4 Branch Orientation

We give a standard orientation to the curve $f(x,y) = 0$ by the tangent vector $(-f_y, f_x)$. This orientation is not intrinsic in the sense that $-f(x,y) = 0$ is the same curve but with opposite standard direction. Given a consistent traversal direction of a branch, we observe that the standard direction of the curve may reverse at certain singularities. For example, the orientation of $f = y^2 - x^2 - x^3 = 0$ is as shown in Figure 5.3. Consequently, when traversing the curve from p to q, we first move in the standard direction, but after the singularity we move in the opposite direction. Figure 5.4 shows that this reversal does not happen at all singularities.

Geometrically, the apparent orientation reversal is understood when considering $f(x,y) = 0$ as the intersection curve of $f(x,y) - z = 0$ and $z = 0$. Along the intersection curve, the projection onto the x-y plane of the surface gradient $(f_x, f_y, -1)$ is just the curve normal. Thus the normal reversal that causes the changed standard orientation is merely a rotation of the surface normal in 3-space. Whether a branch suffers this reversal depends on the global topology of the singularity. Briefly, the orientation reverses if the branch intersects an odd number of other branches, with the proper definition of intersection. Nevertheless, a local correspondence between the orientation of f and its proper transform g can be established outside the singularity from which we can deduce whether the standard orientation has reversed, without having to analyze the topology of the singularity.

Let $p = (a_0, b_0)$ be a nonsingular point of f, where $a_0 \neq 0$. To p corresponds the point $p_1 = (a_0, b_0/a_0)$ of the transformed curve g. Centered at

p, the curve f has the place

$$\begin{aligned} x(s) &= a_0 + a_1 s + a_2 s^2 + \cdots \\ y(s) &= b_0 + b_1 s + b_2 s^2 + \cdots \end{aligned}$$

and centered at p_1, the curve g has the place

$$\begin{aligned} x_1(s) = x(s) &= a_0 + a_1 s + a_2 s^2 + \cdots \\ y_1(s) &= c_0 + c_1 s + c_2 s^2 + \cdots \end{aligned}$$

We assume that the traversal of f at p proceeds by increasing value of s. Note that the traversal direction need not agree with the standard orientation $(-f_y, f_x)$. Since $x(s) = x_1(s)$, the curve and its transform are oriented the same way, and traversing g by increasing s is equivalent to traversing f by increasing s.

Since $y_1(s) = y(s)/x(s)$, we divide the two power series and compare the resulting coefficients with the c_k. We obtain

$$\begin{aligned} c_0 &= b_0/a_0 \\ c_1 &= \frac{b_1 a_0 - a_1 b_0}{a_0^2} \\ &\vdots \end{aligned} \tag{7}$$

Since f is not singular at p, g is not singular at p_1. Hence both curves have a Taylor series at these points so that a_1 is proportional to $-f_y$ and to $-g_y$, while b_1 is proportional to f_x, and c_1 is proportional to g_x. We now obtain from (7) that

$$\begin{aligned} g_y &= \alpha f_y \\ g_x &= \alpha \frac{x f_x + y f_y}{x^2} \end{aligned}$$

So, if we relate the traversal direction of f to the standard orientation $(-f_y, f_x)$, then α relates the corresponding traversal direction of g to the standard orientation $(-g_y, g_x)$ of g, and vice versa. Since the fully desingularized branch cannot experience an orientation reversal, we have a method to maintain consistent traversal direction through singularities.

5.5 Implementation

The planar curve traversal has been implemented in Lisp on a Symbolics Lisp machine. A prototype was previously implemented by C. Bajaj on a VAX 8600. Figure 5.4 shows an example of a traversal requiring iterated desingularization, as well as the traversals along the proper transforms in the vicinity of the singularity. For simplicity, the singularity was already positioned at the origin.

6 Singularities on Space Curves

Every algebraic space curve is the birational image of an algebraic plane curve. It follows that the singularities of space curves are not qualitatively different from those of plane curves. Two possibilities exist for tracing through space curve singularities:

1. Construct a birational map from the given space curve to a planar curve, trace the planar curve, and lift the resulting points.

2. Desingularize the space curve directly.

Since the intersection curve in general has degree equal to the product of the surface degrees, the birationally equivalent plane curve must have high degree and may be computationally less tractable. Working with the space curve directly is therefore more attractive. However, desingularizing the space curve directly must address the fact that the curve is given as the intersection of two surfaces. If we are to work with this representation, then we need to apply quadratic transformations that desingularize the curve and the intersecting surfaces as well. Hence the approach to space curve desingularization found in the standard literature, e.g. [5], will not work, and more research is required to work out techniques suitable for computation. We restrict our discussion therefore to the method of reducing the space curve to a plane curve.

The simplest way to map the space curve to a planar curve is by projection. Orthographic projection along a principal axis is done by elimination of a variable, using resultants. Other directions require a rotation of the coordinate system prior to projection. For space curve singularities where at least one of the surfaces has a nonzero gradient, orthographic projection onto the tangent plane is conceptually ideal. The major computational

problem would be the inefficiency of the resultant computation for surfaces of high degree. Moreover, branches intersecting the surface normal above or below the tangent plane are also projected and unnecessarily increase the complexity of the singularity.

A different method to map a space curve to the plane is to find a rational surface containing the curve, parameterizing this surface, and then substituting the parametric equations into one of the implicit surface equations, say g. This method has the advantage of treating all singularities, not only those at a nonzero surface gradient. We describe the method below and give several examples. It has not been implemented yet.

6.1 Monoid and Cone Representation

It is well known that every algebraic space curve $f \cap g$ can be represented as the intersection of a monoid and a cone, [13]. A *monoid* is a rational surface of degree m that contains an $m-1$ fold point. Simple examples include all planes, quadrics, and the Steiner surface. When the $m-1$ fold point is brought to the origin, the monoid equation takes the form

$$wH_{m-1}(x,y,z) + H_m(x,y,z) = 0$$

where H_{m-1} is homogeneous of degree $m-1$ and H_m is homogeneous of degree m.

We will be interested in determining the monoid containing a given space curve $f \cap g$ and its parametric representation. We will not determine the cone, since it is not needed. Moreover, the parameterization of the monoid is incompatible with the cone equation. The procedure for determining the monoid is based on the projective form of the surfaces f and g. As in [12], we proceed as follows.

First, homogenize $f(x,y,z)$ and $g(x,y,z)$ so as to obtain $F(w,x,y,z)$ and $G(w,x,y,z)$. As long as $w \neq 0$, the curve $F \cap G$ is identical to $f \cap g$. We select one of the base points of the projective coordinate system, say $(1,0,0,0)$, as the $m-1$ fold monoid point. This implies using w as the main variable in the computation below. The base point $(0,1,0,0)$ would correspond to selecting x as the main variable, and so on.

We write both F and G as polynomials in the main variable, w,

$$\begin{aligned} F &= u_n w^n + u_{n-1} w^{n-1} + \cdots + u_1 w + u_0 \\ G &= v_{n'} w^{n'} + v_{n'-1} w^{n'-1} + \cdots + v_1 w + v_0 \end{aligned}$$

Without loss of generality we assume that $n \geq n' > 1$. We compute the polynomials

$$\begin{aligned} F_1 &= u_n w^{n-n'} G - v_{n'} F \\ G_1 &= (u_0 G - v_0 F)/w \end{aligned}$$

Note that both F_1 and G_1 contain the intersection curve of F and G.

Both F_1 and G_1 have degree at most $n-1$ in w. If one of them is linear in w, then we stop; we have found the monoid equation. If neither is linear, then we repeat the calculation using F_1 and G_1 in place of F and G. Since at each step the maximum degree in w is lowered at least by one, the computation derives the monoid equation after at most n steps in the form

$$wH_{m-1}(x,y,z) + H_m(x,y,z) = 0$$

The monoid is parameterized by intersecting it with lines through the $m-1$ fold point. Let $a:b:c$ be the direction ratios of these lines, then the monoid is parameterized by

$$\begin{aligned} w(a,b,c) &= -H_{m-1}(a,b,c)/H_m(a,b,c) \\ x(a,b,c) &= a \\ y(a,b,c) &= b \\ z(a,b,c) &= c \end{aligned}$$

This parameterization is *projective*, that is, (a,b,c) are the coordinates of a two-dimensional projective parameter space.

The parametric forms are now substituted into the equation of G and give the desired plane curve, in homogeneous form.

6.2 Examples

We illustrate the method with two examples. First, consider the intersection curve of the cylinder $F = x^2 + z^2 + 2zw = 0$ and the sphere $G = x^2 + y^2 + z^2 + 4zw = 0$. The intersection curve is an irreducible degree 4 space curve with a nodal singularity at the origin shown in Figure 6.1.

The cylinder is a monoid with the $m-1$ fold point at the origin $(1,0,0,0)$. Since the point of interest on the space curve is the origin, we determine a different monoid whose $m-1$ fold point is not the origin. We choose $(0,0,0,1)$, making z the main variable. Accordingly, we compute

$$F_1 = G - F = y^2 + 2zw$$

The parameterization of F_1 is then

$$\begin{aligned} z &= -2a/c^2 \\ w &= a \\ x &= b \\ y &= c \end{aligned}$$

Substitution into G yields the plane curve

$$b^4 + 4a^2(c^2 - b^2) = 0$$

Dehomogenizing with $a = 1$ yields $b^4 - 4(c^2 - b^2) = 0$. This curve is shown in Figure 6.2.

As a more complicated example, consider the intersection of the torus $(x^2 + y^2 + z^2 - w^2)^2 + 8w^2(z^2 - x^2 - y^2 - w^2) + 16w^4 = 0$ with the ellipsoid $36(z - w)^2 + 4(y - w)^2 + 9x^2 - 36w^2 = 0$. The monoid computation, as described, yields a surface of degree 12 in 4 steps. Its equation contains an extraneous factor of degree 4. Substitution into the ellipsoid equation thus yields a plane curve of degree 16 that factors into a degree 2 component, a degree 6 component, and a degree 8 component.

A better solution to the problem is to parameterize the ellipsoid since it is a monoid. To do so, we first translate the coordinate system to the point $(1, 0, 1, 2)$, which is on the ellipsoid, and parameterize with w as main variable. This results in a plane curve of degree 12 that could not be factored by Macsyma.

6.3 Remarks

Only in the plane curve case do we know of a simple algebraic procedure achieving complete desingularization at minimal computational cost. The projection of space curves to planar curves seems to require significant machinery, for determining the monoid equation, and for eliminating extraneous components introduced in the process. In the case of monoid/cone intersection these extraneous components are all lines, hence would be simple to exclude. The cone is determined by a resultant computation, and yields a homogeneous polynomial in three variables. Unfortunately, its equation is unsuitable for substituting the monoid parameterization. In simple cases, there exist certain reparameterizations that circumvent this problem.

If one of the surfaces f or g is known to be rational it can be advantageous to parameterize it directly. In the case of quadrics this amounts to a coordinate translation that brings one of the base points of the coordinate system onto the surface. In the case of cubics, [2] gives parameterization algorithms. Many other surfaces, including the torus, are also rational with known standard parameterizations. Direct parameterization side-steps a potentially lengthy monoid derivation. It should be noted, however, that the resulting plane curve is not necessarily of minimum degree.

It appears that there are simple quadratic transformations of space that achieve surface/surface intersection desingularization much as in the planar case. More research is needed to explore the exact effect of these transformations, and the issues involved in realizing them by computation.

Acknowledgements

Much of this research was done jointly with C. Bajaj, J. Hopcroft, and R. Lynch.

7 References

1. S. Abhyankar (1983), Desingularization of Plane Curves, *Proc. Symposia in Pure Math. 40*, Part 1, 1–45.

2. S. Abhyankar, C. Bajaj (1987), Automatic Rational Parameterization of Curves and Surfaces II: Cubics and Cubicoids, *Comp. Aided Design*, to appear.

3. C. Bajaj, C. Hoffmann, J. Hopcroft (1987), Tracing Algebraic Curves: Plane Curves, Tech. Rep. CSD-TR-637, Purdue University.

4. G. Collins (1975), Quantifier elimination for real closed fields by cylindrical algebraic decomposition, *2nd GI Conf. on Aut. Thy. and Formal Lang.*, Springer Lect. Notes in Comp. Sci. 33, 134–183. *J. ACM 18*, 515–532.

5. A. Emch (1928) Reduction of Singularities of Space Curves and Surfaces, in *Algebraic Geometry*, V. Snyder et al., eds., Chelsea Publications 1970, 252–256.

6. R. T. Farouki (1986), The Characterization of Parametric Surface Sections, *Comp. Vision, Graphics and Image Proc. 33*, 209–236.

7. R. T. Farouki, V. T. Rajan (1987), On the Numerical Condition of Bernstein Polynomials, *IBM Res. Rept. RC 12626.*

8. C. Hoffmann, R. Lynch (1987), Following Space Curves Numerically, Tech. Rept. CSD-TR-684, Purdue University.

9. M. Pratt, A. Geisow (1986), Surface/surface intersection problems, in *The Mathematics of Surfaces*, J. A. Gregory, ed., Clarendon Press, Oxford, 117–142.

10. J. G. Semple, G. T. Kneebone (1959), *Algebraic Curves*, Clarendon Press, Oxford.

11. J. Schwartz, M. Sharir (1987), On the Piano Movers' Problem II, in *Planning, Geometry, and Complexity of Robot Motion*, J. Schwartz, M. Sharir, J. Hopcroft, ed., Ablex Publishing, Norwood, N.J., 51–96.

12. V. Snyder, C. H. Sisam (1914), *Analytic Geometry of Space*, H. Holt and Co., New York.

13. B. L. van der Waerden (1939), *Einführung in die algebraische Geometrie*, 2nd edition, Springer Verlag, 1973.

14. R. J. Walker (1950), *Algebraic Curves*, Springer Verlag.

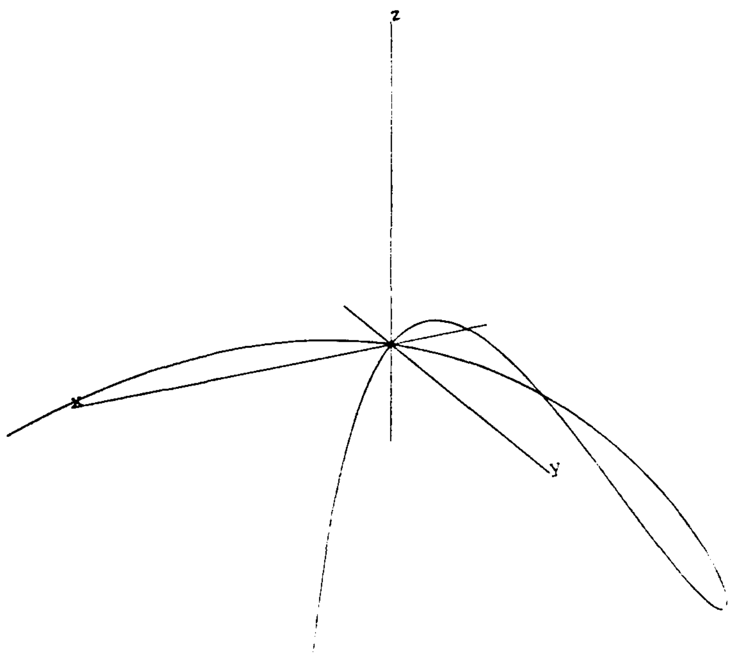


Figure 3.1

$z + y^2 - x^3 \cap z + x^2$
Normal Crossing

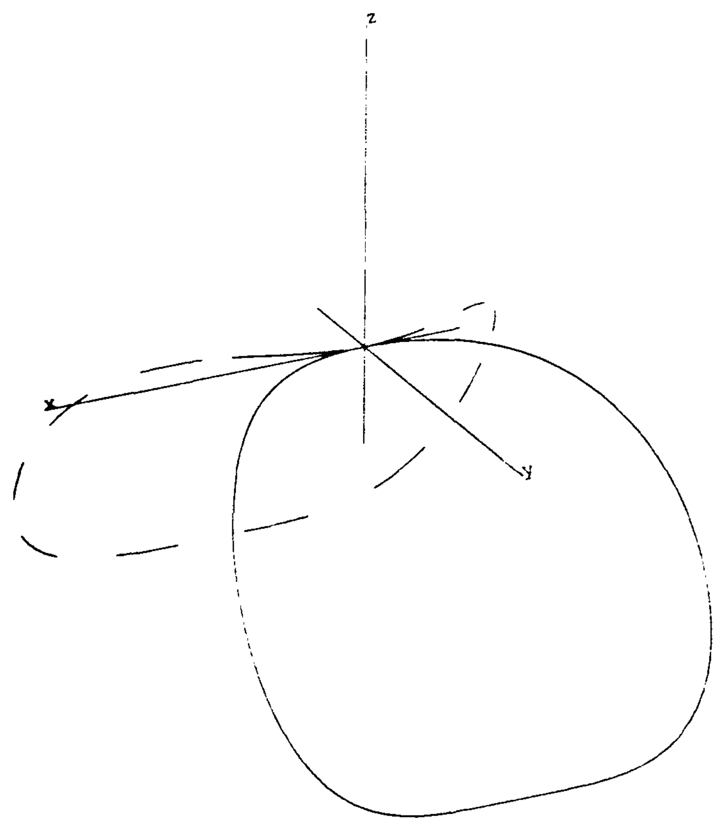


Figure 3.2

$z + x^4 + y^4 \cap z + y^2$
Tacnode

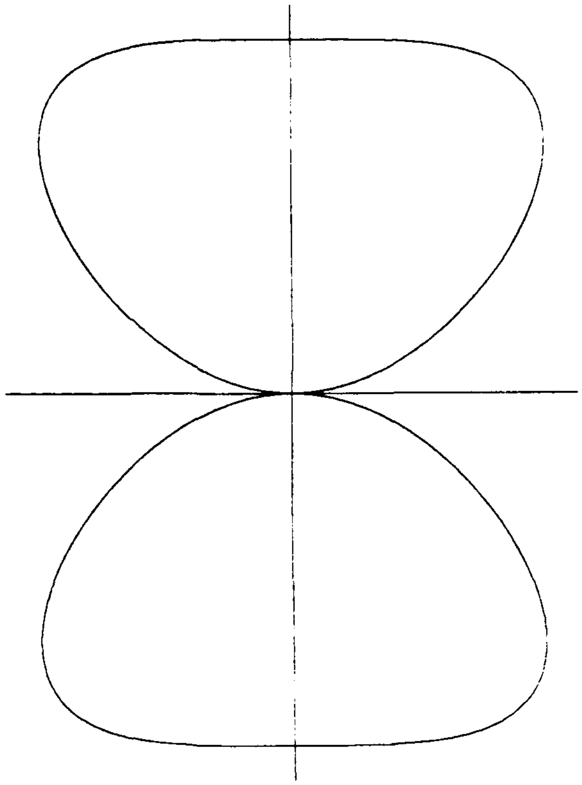


Figure 3.3

$y^2 - x^4 - y^4$
Touching Branches

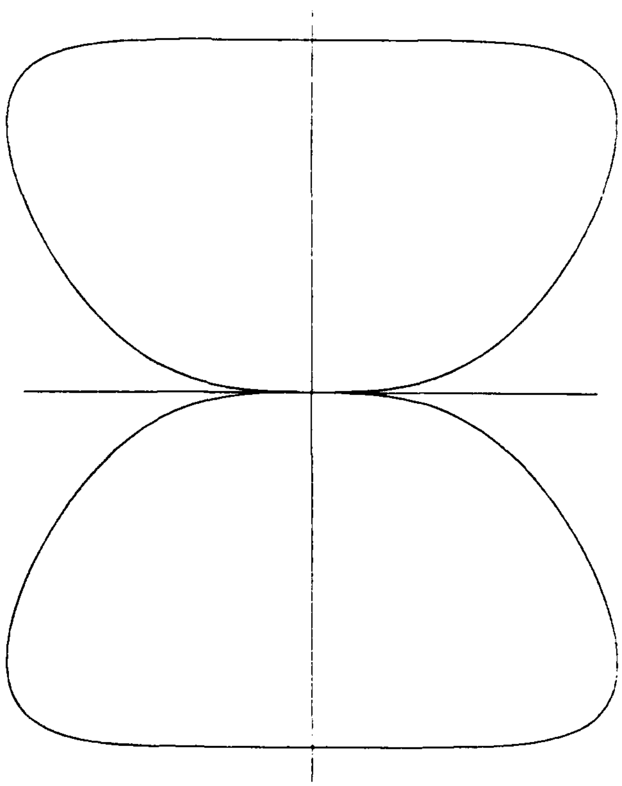
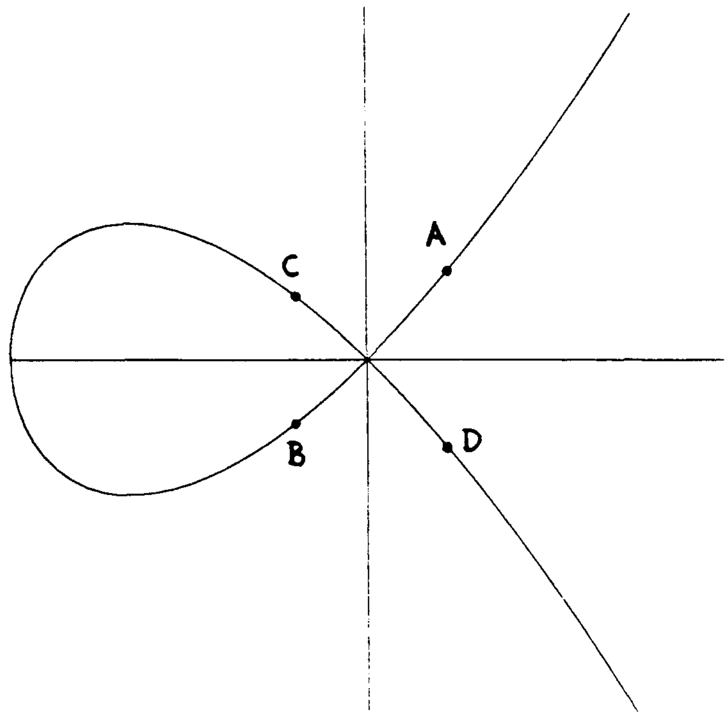


Figure 3.4

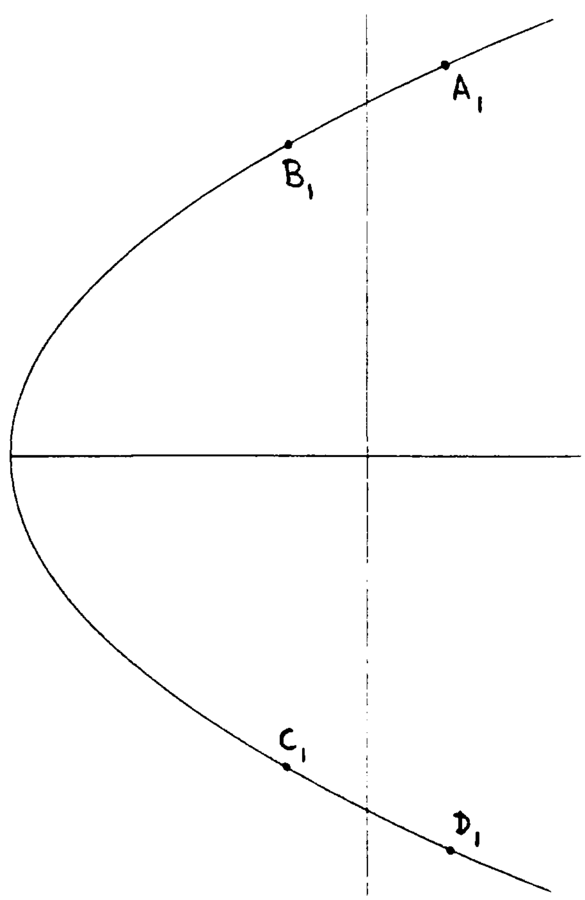
$y^2 - x^6 - y^6$
Intersecting Branches

Figure 5.1

Desingularization of Nodal Singularity



$y^2 - x^2 - x^3$



$y_1^2 - 1 - x_1$

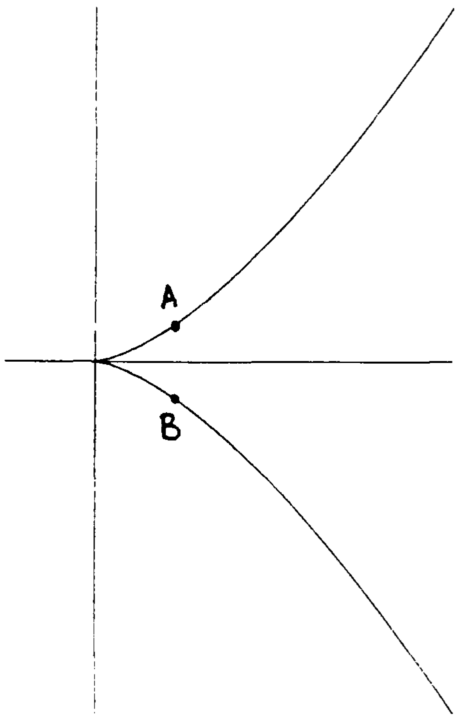
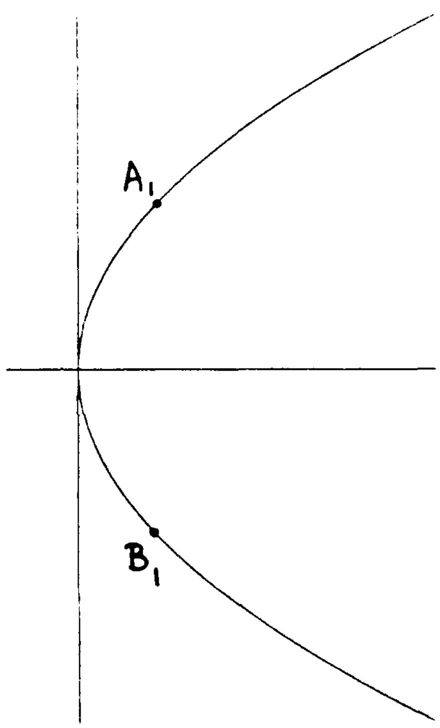


Figure 5.2

Desingularization of Cuspidal Singularity

$y^2 - x^3$



$y_1^2 - x_1$

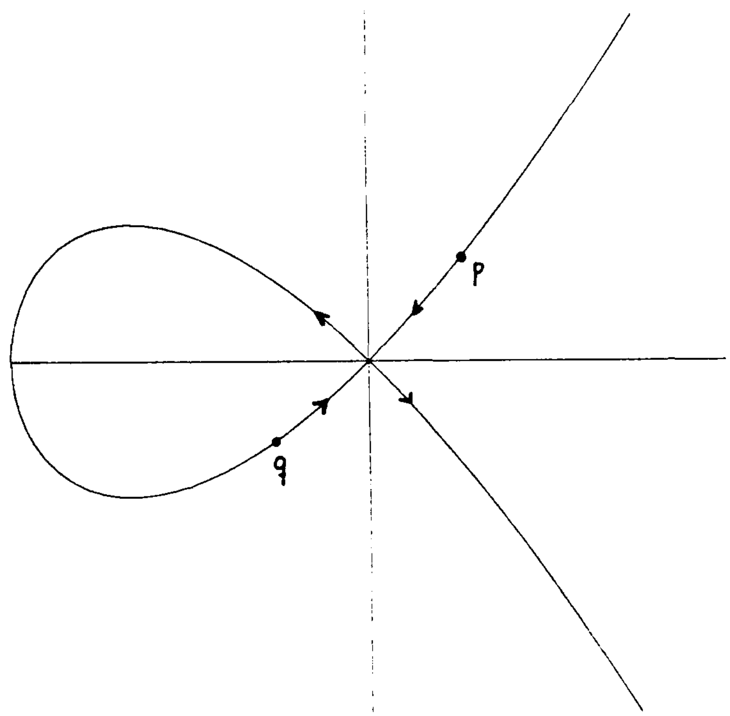


Figure 5.3

$y^2 - x^2 - x^3$
Orientation Reversal at Singularity

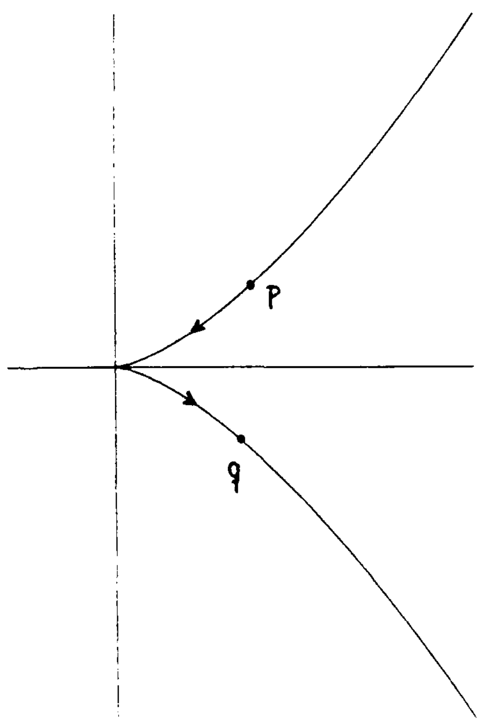


Figure 5.4

$y^2 - x^3$
No Orientation Reversal at Singularity

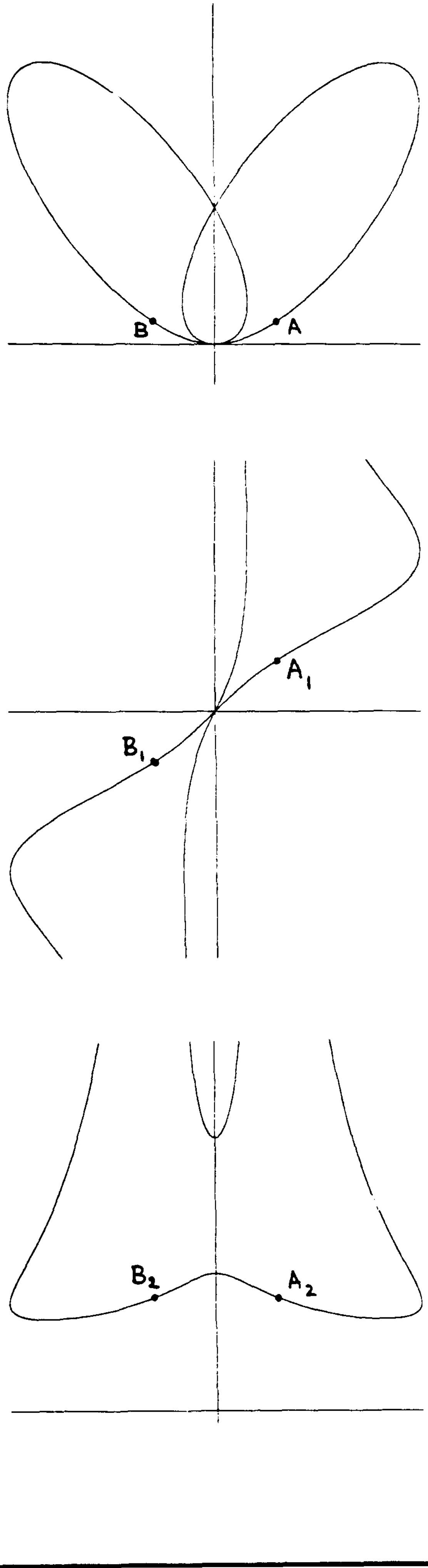


Figure 5.5

Recursive Desingularization

$$2x^4 - 3x^2y + y^2 - 2y^3 + y^4$$

$$2x_1^2 - 3y_1x_1 + y_1^2 - 2y_1^3x_1 + y_1^4x_1^2$$

$$2 - 3y_2 + y_2^2 - 2y_2^3x_2^2 + y_2^4x_2^4$$

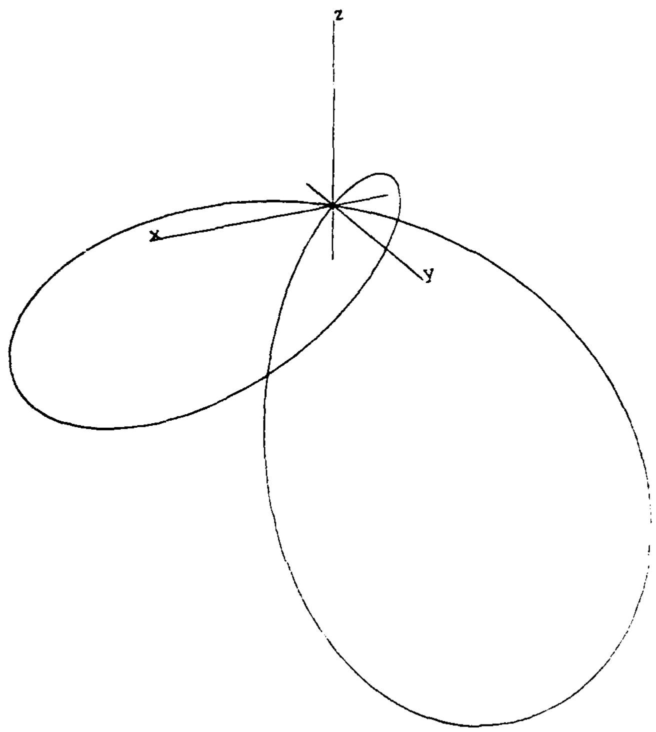


Figure 6.1

Cylinder - Sphere Intersection
$x^2 + y^2 + z^2 + 4z \cap x^2 + z^2 + 2z$

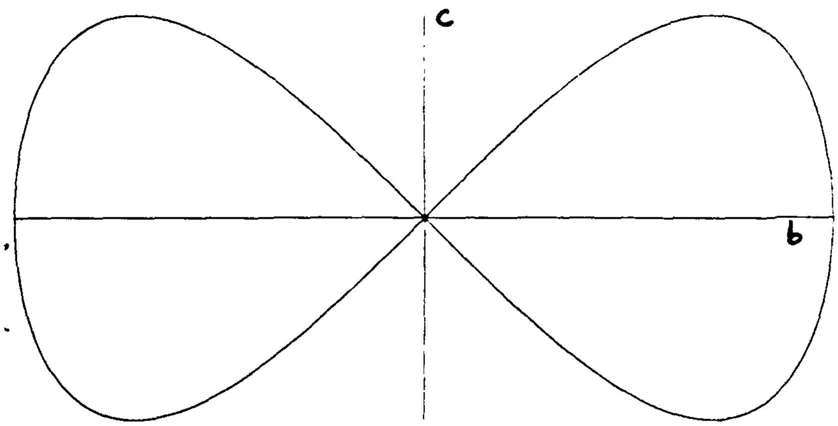


Figure 6.2

Corresponding Plane Curve
$b^4 - 4(c^2 - b^2)$